MODELING THE EFFECTS OF UNCERTAINTY AND RELIABILITY ON THE COST OF ENERGY FROM PV SYSTEMS

Alan Ristow, Miroslav Begović, and Ajeet Rohatgi

School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332 USA

ABSTRACT: The cost of energy produced by a photovoltaic system is dependent upon the amount of energy produced by the system and the amortized cost of the system's components. Existing simulation tools either use crude estimators of system reliability or ignore the effects of system downtime on energy production altogether. Furthermore, the costs associated with system components are often not known precisely during system planning. However, it is difficult to reflect this uncertainty in energy cost calculations using conventional deterministic techniques. This work attempts to address these deficiencies by applying a stochastic model of system reliability to the prediction of energy production over a system's life. Similarly, it uses a stochastic model that encompasses the uncertainties associated with system component prices to estimate the uncertainty in the total installed system cost. Finally, using these two results, it computes the uncertainty in the cost of energy produced by the system. Preliminary testing of this approach, using failure data obtained from an actual system, produces cost estimates of \$0.300–0.400/kWh, with a mean of \$0.349/kWh, consistent with contemporary residential system cost analyses. The link between engineering and economics suggests that the proposed method may be useful as an optimization tool if an appropriate database can be developed from which to draw realistic input distributions.

Keywords: Modeling, Reliability, Economic Analysis

1 INTRODUCTION

The cost of energy produced by a photovoltaic system is dependent upon the amount of energy produced by the system and the amortized cost of the system's components. A number of simulation tools exist for predicting the energy output of a system, fully accounting for system geometry and geography. However, these tools either use crude estimators of system reliability or ignore the effects of system downtime altogether. Furthermore, the cost of the system and its financing is assumed to be well known. In reality, the costs of the system components are often known only approximately until they are actually purchased; this seems to be especially true for installation costs, which may sometimes not be known until after the system is installed. As a result, system planning may be compromised by uncertainty about the final energy cost.

This work attempts to address these deficiencies in cost analysis by applying a stochastic model of system reliability to the prediction of energy production over a given system's life, using a Monte Carlo model to predict the occurrence and duration of system failures. It takes a similar stochastic approach to estimating the system cost, reflecting uncertainties in the costs of system components in the final cost of the installed system. The combination of these calculations yields a probabilistic estimate of the final cost of energy produced by the system, rather than a fixed value that bears no information about its accuracy or sensitivity to downtime or system component costs.

The advantages to this approach are manifold. The output distribution yields an *expected* cost per unit energy and a strongly bounded range of *possible* values for the cost. Furthermore, the distribution makes the probability of meeting or exceeding a given cost target readily apparent. These properties of the stochastic approach give the analyst valuable tools for system planning and analyzing financial risk.

Another advantage of the stochastic approach is that by comparing the output distribution to the distributions of the inputs, the inputs to which the energy cost is most sensitive may be determined. This information may be used to determine how to allocate cost-reduction efforts (for example, the energy cost may be more sensitive to inverter failure rate than to inverter cost). It might also be used prior to purchasing the system in order to seek a more certain cost estimate for a given system component.

Preliminary testing of this approach, using failure data obtained from an actual system, produces cost estimates that are consistent with contemporary residential system cost analyses. The link between engineering and economics suggests that the proposed method may eventually be useful as an optimization tool. In combination with prior work, it may even be able to show the impact of design decisions at the system component level on the final electricity cost, providing a tool to help guide future low-cost PV research.

The Monte Carlo approach has been used previously to investigate the impact of "green power" pricing initiatives [1] and uncertainty in PV module manufacturing cost [2]. This work applies the method to the problem of estimating PV energy production costs. Using a PV system model developed at Georgia Tech [3] and a simple system cost model, this work simulates the effects of random system failures and uncertainties in system component costs on the final energy production costs for a sample PV system. It demonstrates how the power of Monte Carlo simulation may be harnessed to calculate an expected energy cost with a strongly bounded range of possible values, as well as a rigorous sensitivity analysis that traces the uncertainty in the output to specific input variables.

2 PV SYSTEM SIMULATION

The PV system simulator was developed at Georgia Tech, but is based on the well established model PVFORM [4]. The Georgia Tech model calculates module temperatures more accurately and contains provisions for simulating a wider range of system geometries. The model has been validated using data from an operating PV system [3].

The system simulated in this work is a hypothetical south-facing, grid-connected 3 kW_p system located in Atlanta, Georgia. The TMY2 database [5] is used as a climate model to determine the power output from the PV modules at one-hour intervals over the course of a full year. An inverter model determines the AC power output of the system at each interval. Finally, a stochastic reliability model determines the frequency and duration of system failures. Program flow is illustrated in Figure 1. It should be noted that the TMY2 database consists of well characterized, statistically filtered climate observations; therefore, modeling variations in weather conditions was unnecessary. However, because TMY2 is statistically derived, it would be a poor choice of model for investigating the effects of weather extremes [6].

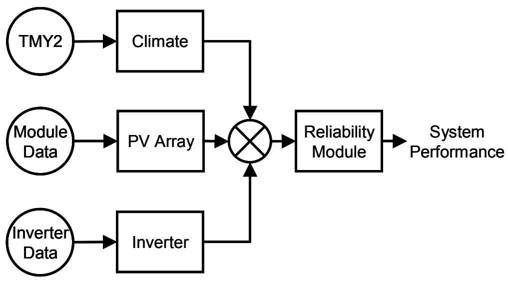

Figure 1: Simulation flow chart.

System failures are modeled using a reliability model consisting of two random functions, one to determine the time to the next failure and another to determine the duration of the failure. The two-parameter Weibull distribution is used to represent these functions because of its ability to take on a wide range of characteristics, including those of other distributions. The probability P of system failure at a time Δt since recovery from the previous failure is calculated from the cumulative distribution function (cdf) for the Weibull distribution,

$$P = 1 - e^{-(\Delta t/\eta)^{\beta}} . \tag{1}$$

For this simulation, the distribution parameters are derived from field data collected over a five-year period from the 342 kW_p PV system atop the Georgia Tech Aquatic Center. The time to failure is determined using location and shape parameters η_1 = 13,213 hr and β_1 = 1.052, respectively, while the duration of failure is determined using η_2 = 718.4 hr and β_2 = 1.7397 [3].

Equation (1) clearly indicates that the probability of system failure increases with Δt. The probability of failure at a particular hour is calculated using η_1 and β_1 and compared to a randomly generated number to resolve whether or not the system fails. If the system does not fail, it continues producing energy (as determined by the system simulation) for that hour; if it fails, it produces no output until the system recovers. The recovery time is computed in the same manner, but uses η_2 and β_2 as parameters. This cycle repeats to the end of system's service life. Because of this random component, 30-year energy production varies from simulation to simulation.

The simulation assumes a 30-year system lifetime with random failures and downtime as described above. The reliability module is applied to the simulated hourly output of the system. The time to the first failure is determined using the first Weibull distribution, and the duration of the failure is determined using the second. During each hour the system is down, its power output is reduced to zero. At the end of the failure the time to the next one is determined by the first Weibull distribution, and the failure's duration by the second. This cycle repeats until the entire 30-year lifespan of the system has been covered, constituting a single sample within the simulation. As a result, each sample will have a different pattern of failures and produce a different amount of energy over its life.

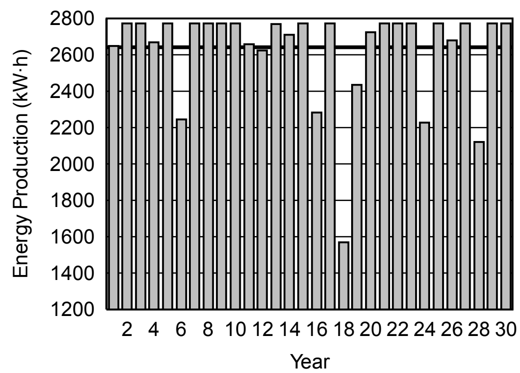

Figure 2: Variation in energy production due to system failures for one sample. The heavy line represents the average annual output for the entire simulation.

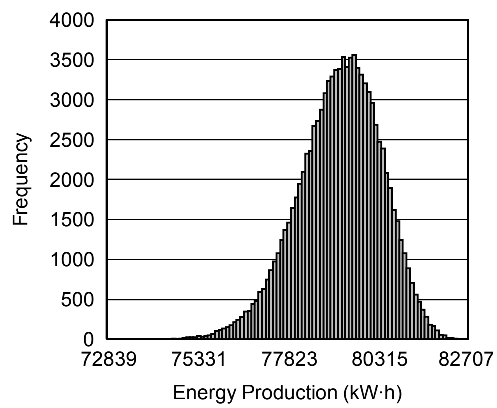

Figure 3: Histogram of lifetime energy production.

Figure 2 shows the simulated energy output of the system over its life for a single representative sample, while Figure 3 illustrates the simulated distribution of energy production over the systems lifetime. The total energy produced by the system over its life is simply the sum of the energy production in each of the system's 30 years. In order to obtain an adequate number of samples for analysis, the simulation must be repeated a sufficient number of times to ensure stability of the statistical moments of the output quantities. The required number of samples is a function of model complexity and input variable specifications. In general, the greater the number of inputs modeled using Monte Carlo, the greater the number of samples required before the model stabilizes. The simulation illustrated here stabilized after about 10,000 samples, as shown in Figure 4.

In this simulation, the mean energy output of the system over 30 years is 79,217 kWh, with a 90%

confidence interval of 77,224–80,961 kWh. The distribution is slightly skewed to the left. The annual energy output, shown in Figure 5, is skewed much more strongly to the left and appears to be bimodal, with a strong peak at 2760 kWh and a weak peak 2718 kWh. It has a mean of 2641 kWh and a 90% confidence interval of 2176–2771 kWh. The frequency at the strong peak is just over 1.5x106, representing about half of the years included in the 100,000 simulations (which each span 30 years), and the value of this peak is 2760 kWh, equal to a year with no failures. This indicates that about 15 years of each 30-year simulation pass without a single failure. In combination with Figure 3, these data indicate that the effects of system failures on 30-year energy production are moderated by time and the random nature of the failures. While individual years may see greatly reduced production, the amount of energy lost to downtime over the system's life is unlikely to be more than 7%, based on the lower limit of the 90% confidence interval for 30-year energy production.

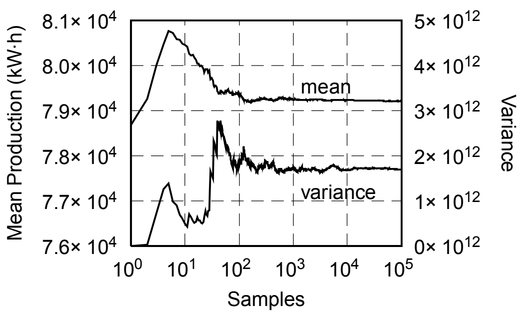

Figure 4: Mean and variance of the total energy production as a function of the number of Monte Carlo samples, showing stability after about 10,000 samples.

The cost of the system C, in dollars per watt, is calculated from

$$C = P_{\text{mod}} + P_{\text{bos}} + P_{\text{ins}} + \frac{A_{\text{bos}}}{1000\eta}, \tag{2}$$

where P_{mod} is the module cost ($/W$_p$), P_{bos} is the power-related BOS cost ($/W$_p$), P_{ins} is the installation cost ($/W$_p$), A_{bos} is the area-related BOS cost ($/m^2), and η is the nominal module efficiency (10.9% in this simulation). The cost is amortized over the 30-year life of the system using standard amortization equations, with simple deductions for U.S. federal tax credits on interest payments and adjustments for inflation [7]. The inflation rate was assumed to be 4.6%, the average inflation rate in the United States over the past 30 years [8].

If only the minimum and maximum possible values for each variable were known, the uniform distributions could be assumed and the system cost could be computed using simple interval arithmetic instead of stochastic methods. However, when the values in the range are not equally probable a method such as Monte Carlo is required. For this simulation, the model calculates the system cost using triangular distributions for the costs of system components using the parameters shown in Table I. For each of the samples, the simulated system cost is divided by the energy produced to yield an estimate of the system cost per unit energy. The aggregate of these samples may be used to produce a probability distribution for the system's energy cost.

Table I: Parameters for triangular distribution of system component costs and financial parameters.

	Minimum	Most Probable	Maximum
Module cost ($/W$_p$)	3.00	4.00	5.00
Power BOS ($/W$_p$)	0.50	0.90	1.30
Installation ($/W$_p$)	1.00	1.50	2.00
Area BOS ($/m^2)	75.00	125.00	175.00
Interest rate (APR)	6.0	7.5	9.0

It is important to note that both the distribution and the parameter values critically affect the results of the simulation. For application to a real PV system, input distributions must be carefully selected based upon available data relevant to the project if the modeled output distributions are to be meaningful.

3 RESULTS AND ANALYSIS

The probability density function for the energy cost is shown in Figure 6. The mean cost of the electricity produced by the system is $0.349/kWh, with a 90% confidence interval of $0.300–0.400/kWh. Thus, while the system is expected to produce electricity at $0.349/kWh, the actual cost can lie anywhere within the confidence interval. The true cost will depend on the actual cost of the system components, the actual interest rate paid, and the actual incidence of system failures and their durations. In contrast, the energy cost calculated in the traditional manner, using only the most probable values from Table I and assuming the system never fails, is $0.332/kWh. While this figure is close to the mean predicted by the Monte Carlo simulation, it provides no additional information by which to judge its accuracy.

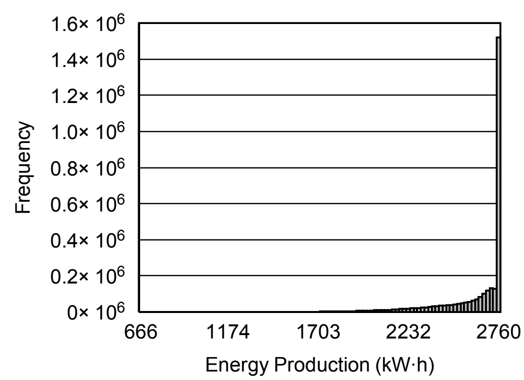

Figure 5: Histogram of annual energy production.

The output distribution may provide valuable information about the viability of a design such as this one, given that the input assumptions are correct. It may also be used to re-evaluate a particular system as component prices become known with greater certainty. For example, just prior to construction, all system components have been purchased and their costs are known exactly, except perhaps the installation cost. Likewise, financing will have been obtained and the interest rate will also be known. The incidence of failures, of course, will not be known until the end of the system's life. By running the model again, one may obtain a more precise estimate of the energy cost of the system. Thus, one may use this tool to analyze either a class of system designs or a specific system.

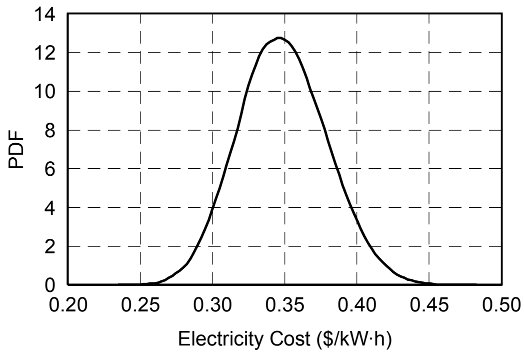

Figure 6: Probability density function for the cost of energy produced over 30 years. Mean $0.349/kWh, 90% confidence interval $0.300–0.400/kWh.

It is important to note that obtaining meaningful output distributions from this model is critically dependent upon having a large database of metrics from existing systems from which to extract accurate input distributions. To this end, it is desirable to maximize the number of systems included in the database, and to monitor those systems over the longest time intervals possible. As a result, the accuracy of the proposed model might reasonably be expected to improve over time as the set of available input data expands.

Figure 7: Sensitivity of energy cost to the input parameters. The standardized regression coefficient indicates the change in energy cost for the same relative change in any of the input variables.

Sensitivity analysis was performed by the method of standardized regression coefficients [9]. The energy cost calculations were regressed against the input variables from Table I and the number of failures in the system's lifetime. The results appear in Figure 7, which shows the simulated uncertainty in energy cost for this system results primarily from the uncertainties in module cost and interest rate. This may be due to high sensitivity of the model to these parameters, or to the range and distribution of the uncertainty in the parameters, or to a combination of the two. In any case, it indicates that the precision of the energy cost estimate may be most expediently increased by increasing the precision of the module cost or interest rate estimates. Additional sensitivity analyses, not performed here, can be used to determine how much of the energy cost uncertainty is due to the distribution of the inputs and how much is due to model sensitivity.

4 CONCLUSIONS

The authors have applied the Monte Carlo method to the problem of simulating PV system reliability and estimating the cost of electricity produced by a grid-connected PV system. The stochastic nature of the model allows it to both use and produce a greater range of information than traditional simulation methods do, and the statistical tools that may be brought to bear on the output data provide powerful methods for gaining insight into system design and behavior. As a result, the approach demonstrated here provides a flexible, multifaceted tool for investigating PV system cost components and behavior.

The stochastic approach illustrated in this work, when applied to multiple system designs, has potential as a decision tool to compare each alternative and select the best design for a given a set of input parameters. Future work will focus on using experimental data to produce input distributions that are highly representative of real PV systems. It will also focus on the modeling techniques required to increase the number of variables included in the simulation. Since the computational expense incurred by Monte Carlo simulation is dependent upon both the probability distributions and the number of random variables in the model, this effort may require investigation of variance reduction techniques.

REFERENCES

[1] M. Begović, A. Pregelj, A. Rohatgi, and C. Honsberg, "Green power: Status and perspectives," *Proc. IEEE* **89**, pp. 1734–1743, Dec. 2001.

[2] A. Ristow, M. Begović, and A. Rohatgi, "Numerical approach to uncertainty and sensitivity analysis in forecasting the manufacturing cost and performance of PV modules," in *Proc. 19th Eur. Photovoltaic Solar Energy Conf.*, (Paris, France), pp. 2178–2181, 7–11 June 2004.

[3] A. Pregelj, *Impact of Distributed Generation on Power Network Operation*. PhD thesis, Georgia Institute of Technology, Atlanta, Georgia, 2003.

[4] D. F. Menicucci and J. P. Fernandez, "User's manual for PVFORM: Photovoltaic system simulation program for standalone and grid-interactive applications," Final Report SAND85-0376, Sandia National Laboratories, 1989.

[5] W. Marion and K. Urban, *User's Manual for TMY2s*. National Renewable Energy Laboratory, Golden, Colorado, June 1995.

[6] A. Pregelj, M. Begović, A. Rohatgi, and D. Novosel, "Renewable distributed resources analysis using quantitative methods for large data sets," *IEEE Trans. Power Sys.*, **19**(3), pp. 1277–1285, 2004.

[7] S. A. Ross, R. W. Westerfield, and B. D. Jordan, *Essentials of Corporate Finance*, Richard D. Irwin, Chicago, 1996.

[8] U.S. Bureau of Labor Statistics, Consumer Price Index, All Urban Consumers (CPI-U), 2004.

[9] R. L. Iman and J. C. Helton, "An investigation of uncertainty and sensitivity analysis techniques for computer models," *Risk Analysis* **8**(1), pp. 71–90, 1988.